

Open Research Online

The Open University's repository of research publications and other research outputs

Not Wanted: On Scharp's Solution to the Liar

Journal Item

How to cite:

Pinder, Mark (2021). *Not Wanted: On Scharp's Solution to the Liar*. *Erkenntnis*, 86 pp. 1567–1584.

For guidance on citations see [FAQs](#).

© 2019 Mark Pinder



<https://creativecommons.org/licenses/by/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.1007/s10670-019-00170-x>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk



Not Wanted: On Scharp's Solution to the Liar

Mark Pinder¹

Received: 20 August 2018 / Accepted: 16 September 2019
© The Author(s) 2019

Abstract

Kevin Scharp argues that the concept of truth is defective, and is therefore unable to play its intended role in natural language truth-conditional semantics. As such, for this theoretical purpose, Scharp constructs two replacements: ascending truth and descending truth. Scharp applies the resultant theory, AD semantics, to the liar sentence, thereby obtaining a novel solution to the liar paradox. The aim of the present paper is fourfold. First, I show that, contrary to Scharp's claims, AD semantics in fact yields an inconsistency when applied to standard liar sentences. Second, I diagnose the problem: AD semantics mishandles negation. I propose an alternative treatment, resulting in what I call AD* semantics. Third, I show that AD* semantics gives Scharp the resources required to respond to an alleged revenge paradox that has been raised against his view. Finally, I argue that, these consequences notwithstanding, it remains unclear whether AD* semantics provides an adequate account of alethic paradoxes more generally.

1 Introduction

Kevin Scharp believes that many of our concepts are defective.¹ He tells us that, sometimes, a concept's defect may prevent it from playing its intended explanatory role in some field of inquiry. In such a case, Scharp recommends replacing the defective concept with something better suited to playing the relevant explanatory role.

In light of the liar and other alethic paradoxes, Scharp believes that our ordinary concept of truth, TRUTH, is defective.² In contrast, he deems the logical principles required to derive the paradoxes unproblematic. Scharp argues that TRUTH's defect prevents it from playing its intended explanatory role in truth-conditional, natural

¹ See e.g. Scharp (2013a: 290).

² See, especially, Scharp (2013a, 2019). See also Scharp (2007, 2013b, 2014).

✉ Mark Pinder
mark.jonathan.pinder@gmail.com

¹ The Open University, Milton Keynes, UK

language semantics: when we try to give truth-conditional semantic theories for fragments of natural languages that contain the resources to construct liar sentences, i.e. sentences of the form:

(1) Sentence (1) is not true,

we end up deriving contradictions. Scharp concludes that, for this theoretical purpose, we should replace TRUTH.

Scharp recommends that we approach the issue with the mindset of a conceptual engineer: we should replace TRUTH with a team of newly engineered concepts, ASCENDING TRUTH and DESCENDING TRUTH. The informal, guiding idea is twofold. Firstly, for each concept, we have only one direction of the unrestricted T-schema. Whenever “X” denotes the sentence replacing “*p*”:

If *p*, then X is ascending true.

If X is descending true, then *p*.

Secondly, the converses of these conditionals hold only for so-called *safe* sentences:

If X is safe, then: if X is ascending true, then *p*.

If X is safe, then: if *p*, then X is descending true.

It follows that a safe sentence is such that its ascending and descending truth values coincide. All other sentences are *unsafe*, and are both ascending true and descending false.³ (Given the principles above, no sentence can be both ascending false and descending true.) The concepts ASCENDING TRUTH and DESCENDING TRUTH are formally introduced by an axiomatic theory, ADT, which I introduce in Sect. 2.

The concepts ASCENDING TRUTH and DESCENDING TRUTH are designed to play the theoretical role traditionally played by TRUTH in truth-conditional, natural language semantics, the dominant approach to the empirical study of linguistic meaning in natural languages.⁴ Scharp develops a variant of the approach, called *AD semantics*. (To emphasise: AD semantics is an approach to the *empirical study of meaning in natural languages*; it is *not* an intended interpretation of the formal theory ADT).⁵ Whereas, for each sentence of the object language, truth-conditional semantic theories yield a biconditional that encodes a *truth* condition for that sentence, AD semantic theories yield *both* a biconditional encoding an *ascending* truth condition and a biconditional encoding a *descending* truth condition.

Developing AD semantics, Scharp provides an assessment-sensitive semantics for the English word “true”.⁶ The underlying idea is that, *qua* linguists, we can choose

³ Throughout, I say that a sentence is *ascending false* just in case it is not ascending true, and *descending false* just in case it is not descending true.

⁴ Scharp emphasises this as the principal motivation for developing ASCENDING TRUTH and DESCENDING TRUTH in his 2019.

⁵ For the intended interpretation of ADT, see Scharp (2013a: 178–187)

⁶ See especially Scharp (2013a, b).

to interpret an ordinary utterance of “true” *either* as expressing ASCENDING TRUTH, *or* as expressing DESCENDING TRUTH. Scharp claims that, either way, sentences such as (1) are found to be unsafe. In addition he claims that, whichever way we interpret an utterance of “true”, the reasoning involved in the liar paradox turns out to be invalid. Likewise for other alethic paradoxes, such as Curry’s paradox and Yablo’s paradox.⁷ These findings constitute Scharp’s solution to the liar and other alethic paradoxes.

Scharp argues that this solution has a major advantage over its competitors: it does not even *face* a revenge paradox.⁸ That is, Scharp argues that the usual strategies for constructing revenge sentences simply do not give rise to anything that the view cannot handle. Moreover, this is achieved without any expressive restrictions, without imposing an object/meta language distinction, and without endorsing a theory of truth that implies that its own axioms are not true. If right, this is an impressive result.

The aim of the present paper is fourfold. First, in Sect. 3, I show that, as it stands, AD semantics in fact does *not* yield the results Scharp claims: surprisingly, it yields an inconsistency even when applied to *standard* liar sentences. It follows that AD semantics is no advance on truth-conditional semantics, and directly undermines Scharp’s solution to the liar, AD semantics, and the principal theoretical motivation for ADT.

Second, in Sect. 4, on behalf of Scharp, I diagnose the problem. I argue that AD semantics mishandles negation, and show how to amend it accordingly. I call the resulting theory AD* semantics.

Third, in Sect. 5, I show that this result has an important consequence for Scharp: it gives him the resources required to respond to an alleged revenge paradox, which I raised in Pinder (2015). Using AD* semantics in place of AD semantics, the derivation of the revenge paradox is blocked, and we instead derive the intuitively correct result that the revenge sentence is *unsafe*.

Finally, in Sect. 6, I argue that, these consequences notwithstanding, it remains unclear whether AD* semantics provides an adequate account of alethic paradoxes more generally. AD* semantics, when combined with ADT, is stronger than ADT alone. The consequence is that, as things stand, there is no guarantee that AD* semantics can be consistently applied to ‘pathological’ sentences in general.

2 ADT and AD Semantics

ADT is a consistent, axiomatic theory expressed in first-order predicate logic (FOL).⁹ ADT consists of a list of twenty axiom schemas; I only introduce part of the theory herein.

⁷ See Scharp (2013a: 255f).

⁸ See Scharp (2013a, especially 271–273).

⁹ For Scharp’s consistency proof, see his (2013a: 178–187). Bacon (2019) also provides a simple model for ADT.

Let \mathcal{L} be a first-order language consisting of: the logical connectives ($\sim, \vee, \wedge, \rightarrow, \leftrightarrow$); constants (a, b, \dots); n -ary predicates (P_n, Q_n, \dots) and functions (f_n, g_n, \dots); variables (x, y, \dots); quantifiers (\exists, \forall); parentheses; three unary predicates (A, D, S); and a device ($\langle \cdot \rangle$) for forming names of sentences of \mathcal{L} . The guiding idea behind ADT is initially captured by three axiom schemas. Where “ ϕ ” can be replaced by any sentence of \mathcal{L} :

- (A-In) $\phi \rightarrow A\langle\phi\rangle$
 (D-Out) $D\langle\phi\rangle \rightarrow \phi$
 (S) $S\langle\phi\rangle \leftrightarrow (D\langle\phi\rangle \vee \sim A\langle\phi\rangle)$.

Each instance of the following is thus a simple theorem of ADT:

- (A-Out) $S\langle\phi\rangle \rightarrow (A\langle\phi\rangle \rightarrow \phi)$
 (D-In) $S\langle\phi\rangle \rightarrow (\phi \rightarrow D\langle\phi\rangle)$

Traditional truth-theoretic rules for connectives, such as:

$$\begin{aligned} T\langle\phi\rangle &\leftrightarrow \sim T\langle\sim\phi\rangle \\ T\langle\phi\wedge\psi\rangle &\leftrightarrow (T\langle\phi\rangle \wedge T\langle\psi\rangle) \end{aligned}$$

do not straightforwardly hold for ASCENDING TRUTH or DESCENDING TRUTH. But ADT gives us various analogues, such as the following axiom schemas:

- (M) $D\langle\phi\rangle \leftrightarrow \sim A\langle\sim\phi\rangle$
 (\wedge -A) $(A\langle\phi\rangle \wedge A\langle\psi\rangle) \rightarrow A\langle\phi\wedge\psi\rangle$
 (\wedge -D) $D\langle\phi\wedge\psi\rangle \rightarrow (D\langle\phi\rangle \wedge D\langle\psi\rangle)$

For the full axiomatisation of ADT, see Scharp (2013a: 154).

It is easy to show that ‘ascending liars’ and ‘descending liars’, i.e. sentences of the form:

$$\begin{aligned} \lambda_A &\sim A\lambda_A \\ \lambda_D &\sim D\lambda_D \end{aligned}$$

where “ λ_A ” and “ λ_D ” are names of $\sim A\lambda_A$ and $\sim D\lambda_D$ respectively, are unsafe. That is, $\sim S\lambda_A$ and $\sim S\lambda_D$ are theorems of ADT. Ascending liars and descending liars are *both* ascending true *and* descending false.

ADT serves to formally introduce ASCENDING TRUTH and DESCENDING TRUTH, which are engineered principally to play the central theoretical role in AD semantics. Recall that AD semantics is an approach to the empirical study of meaning in natural languages, developed to contain an assessment-relative semantics for “true”. AD semantics is divided into three parts, which I will call: an informal *presemantic theory*, an axiomatic *core semantic theory*, and a semi-formal *postsemantic theory*.¹⁰ Together, these theories allow us to derive biconditional ascending/descending truth conditions for utterances in natural language. I will sketch an AD semantic theory

¹⁰ Here, Scharp draws on work by Predelli (2005). See Scharp (2013a: 207–215, 225–253).

for the fragment of English that contains: “sentence (1)” as used in this article, i.e. as a name for the sentence

Sentence (1) is not true;

the predicates “true” and “written”; “not” and “and”; and the copula “is”. Call this fragment *English-*.

A presemantic theory takes a real-world utterance of natural language and provides formal representations of: the uttered sentence; and the relevant features of the context of utterance and assessment. The relevant features are represented by an *index*:

$$\langle w, t, d, s \rangle$$

where w is the world of utterance, t is the time of utterance, d is the agent who performs the utterance, and s is the *alethic standard*. As we see below, an alethic standard encodes how “true” is to be interpreted. According to Scharp, the alethic standard is fixed *not* by the utterer, but by whoever is *using* the semantic theory (i.e. the theorist). As only the alethic standard is relevant in what follows, I omit the other constituents of indexes.

Our presemantic theory for English- will use ε - to represent utterances in English-. ε - contains the following:

- Individual constant: λ .
- One-place predicates: T ; W .
- Logical connectives: \sim ; \wedge .

And the syntactic rules for ε —are as follows:

- For individual constant α and one-place predicate T , “ $T\alpha$ ” is a sentence.
- For sentences ϕ and ψ , “ $\sim\phi$ ” and “ $\phi\wedge\psi$ ” are sentences.

The informal presemantic theory might then be stated as follows:

Utterances in English- are represented in ε -, taking syntactic differences into account as appropriate, where: “ λ ” represents “sentence (1)”; “ T ” represents “is true”; “ W ” represents “is written”; “ \sim ” represents “not”; and “ \wedge ” represents “and”. When the assessor interprets “true” as an ascending truth predicate, the context of assessment is represented as “SA”; when she interprets “true” as a descending truth predicate, the context of assessment is represented as “SD”.

Thus, for example, our presemantic theory will represent “sentence (1) is written” as:

$$W\lambda.$$

A core semantic theory is an axiomatic theory that yields semantic values for sentences of a formal language on the basis of their constituents and modes of composition. Our core semantic theory is formulated in a rich metalanguage of ε - that

contains the language of set theory. As AD semantic theories ultimately yield both *ascending* truth conditions and *descending* truth conditions, we need to distinguish between what we might call *ascending semantic values* (at alethic standard *s*), represented by the function $^A\llbracket \cdot \rrbracket_s$, and *descending semantic values* (at alethic standard *s*), represented by the function $^D\llbracket \cdot \rrbracket_s$.¹¹ As usual, the semantic values of individual constants are objects, the semantic values of predicates are sets of objects, and the semantic values of sentences are 1 or 0.

The core semantic theory for ε - defines the semantic value for λ thus:

$$(I) \quad ^V\llbracket \lambda \rrbracket_s = \text{“sentence (1) is not true”}$$

In AD semantics, the ascending and descending semantic values of a predicate “*I*” can in principle come apart: $^A\llbracket I \rrbracket_s$ is the set of objects that “*I*” is ascending-true of; and $^D\llbracket I \rrbracket_s$ is the set of objects that “*I*” is descending-true of.¹² We can thus define the semantic value for *W* thus:

$$(II) \quad ^A\llbracket W \rrbracket_s = \{o : \text{it's ascending true that } o \text{ is written}\}$$

$$(III) \quad ^D\llbracket W \rrbracket_s = \{o : \text{it's descending true that } o \text{ is written}\}$$

The semantic value for “*T*” is given relative to the alethic standard. The core semantic theory treats “*T*” as expressing: ASCENDING TRUTH at SA; and DESCENDING TRUTH at SD. Formally, Scharp achieves this via a shift in the language.¹³ Let English-A be the language obtained by replacing every instance of “true” in English- with “ascending true”; and let English-D be the language obtained by replacing every instance of “true” in English- with “descending true”. And, if *o* is a sentence of English-, let o_A and o_D be the corresponding sentences of English-A and English-D respectively. Then:

$$(IV) \quad ^A\llbracket T \rrbracket_{SA} = \{o : \text{it's ascending true that } o_A \text{ is ascending true in English-A}\}$$

$$(V) \quad ^A\llbracket T \rrbracket_{SD} = \{o : \text{it's ascending true that } o_D \text{ is descending true in English-D}\}$$

$$(VI) \quad ^D\llbracket T \rrbracket_{SA} = \{o : \text{it's descending true that } o_A \text{ is ascending true in English-A}\}$$

$$(VII) \quad ^D\llbracket T \rrbracket_{SD} = \{o : \text{it's descending true that } o_D \text{ is descending true in English-D}\}$$

Finally, the core semantic theory defines the semantic value of sentences recursively as follows. Where α is an individual constant, *I* a predicate, and ϕ and ψ sentences:

$$(VIII) \quad ^V\llbracket I\alpha \rrbracket_s = 1 \text{ iff } ^V\llbracket \alpha \rrbracket_s \in ^V\llbracket I \rrbracket_s.$$

$$(IX) \quad ^V\llbracket \sim\phi \rrbracket_s = 1 \text{ iff } ^V\llbracket \phi \rrbracket_s = 0.$$

$$(X) \quad ^V\llbracket \phi \wedge \psi \rrbracket_s = 1 \text{ iff } ^V\llbracket \phi \rrbracket_s = 1 \text{ and } ^V\llbracket \psi \rrbracket_s = 1.$$

¹¹ This is my terminology, not Scharp's. Rather than distinguishing ascending semantic values and descending semantic values, Scharp recursively defines truth-in-a-model relative to an *alethic value parameter* which can take *ascending truth* or *descending truth* as its value (2013a: 214–215). Formally, the approaches are mere notational variants, so long as we treat $^A\llbracket \cdot \rrbracket_s$ and $^D\llbracket \cdot \rrbracket_s$ as functions obtained by setting the value of parameter *V* in the function $^V\llbracket \cdot \rrbracket_s$.

¹² For Scharp's discussion of predication, see his (2013a: 218–221).

¹³ See Scharp (2013a: 249).

For example, setting the alethic standard to SA, the core semantic theory yields the following biconditional for “Wλ”:

$$\begin{aligned} {}^A\llbracket W\lambda \rrbracket_{SA} = 1 \quad & \text{iff } {}^A\llbracket \lambda \rrbracket_{SA} \in {}^A\llbracket W \rrbracket_{SA} & [\text{By (VIII)}] \\ & \text{iff “sentence (1) is not true”} \in \{o : \text{it's a-true that } o \text{ is written}\} & [\text{By (I) and (II)}] \end{aligned}$$

Given that, as a matter of empirical fact, “sentence (1) is not true” *is* written, it straightforwardly follows from (A-In) that it is ascending true that that sentence is written.¹⁴ Thus, the right-hand side of the biconditional holds, and we can infer that ${}^A\llbracket W\lambda \rrbracket_{SA} = 1$.

Finally, a postsemantic theory is a semi-formal theory that translates the theorems of the core semantic theory into an ascending and descending truth values for the original utterance. Our postsemantic theory may be stated thus:

An uttered sentence, p , is *ascending true* from a context of assessment iff: where ϕ is the representation of p and s is the representation of the context of assessment, ${}^A\llbracket \phi \rrbracket_s = 1$.

An uttered sentence, p , is *descending true* from a context of assessment iff: where ϕ is the representation of p and s is the representation of the context of assessment, ${}^D\llbracket \phi \rrbracket_s = 1$.

Thus, for each sentence of English-, we obtain both an *ascending* and a *descending* truth condition. Given that ${}^A\llbracket W\lambda \rrbracket_{SA} = 1$, we can conclude that “sentence (1) is written” is ascending true from SA.

Scharp applies AD semantics to a fragment of English containing the liar paradox.¹⁵ To illustrate how he does this, let us apply our theory to (1). The presemantic theory will model (1) as:

$$\sim T\lambda.$$

Scharp supposes (for concreteness) that we have selected the alethic standard SD. Scharp then notes that, in SD, “T” is interpreted as a descending truth predicate, and that descending liars, such as λ_D , are ascending true and descending false. He immediately concludes that:

$$\begin{aligned} (2) \quad (a) \quad & {}^A\llbracket \sim T\lambda \rrbracket_{SD} = 1 \\ (b) \quad & {}^D\llbracket \sim T\lambda \rrbracket_{SD} = 0. \end{aligned} \quad {}^{16}$$

Given (2), the postsemantics straightforwardly tells us that: (1) is ascending true, and not descending true, from our context of assessment. That is, (1) is

¹⁴ See Scharp (2013a: 170).

¹⁵ See Scharp (2013a: 247–256).

¹⁶ See the top of Scharp (2013a: 252)

unsafe—intuitively, the right result. Scharp claims that, had we instead selected the alethic standard SA, “the results would have been the same” (2013a: 252).

3 Inconsistency

There is a striking feature of Scharp’s discussion. Despite developing substantive formal machinery, he only reasons *informally* when explicitly addressing the liar paradox. In particular, Scharp does not *prove* (2a, b), merely relying on an informal comparison to descending liars. Unfortunately, the details of AD semantics do not bear out the result.

As before, let us suppose that the presemantic theory models (1) as “ $\sim T\lambda$ ” and sets the alethic standard to SD. Now, consider what the semantic theory tells us about $\sim T\lambda$:

$$(3) \quad \begin{array}{lll} \nu\llbracket \sim T\lambda \rrbracket_{SD} = 1 & \text{iff } \nu\llbracket T\lambda \rrbracket_{SD} = 0 & [\text{By (IX)}] \\ & \text{iff } \nu\llbracket \lambda \rrbracket_{SD} \notin \nu\llbracket T \rrbracket_{SD} & [\text{By (VIII)}] \end{array}$$

Focusing on the *descending* semantic value, we thus have:

$$D\llbracket \sim T\lambda \rrbracket_{SD} = 1 \text{ iff } D\llbracket \lambda \rrbracket_{SD} \notin D\llbracket T \rrbracket_{SD}$$

From (I) and (VII), we can thus infer:

$$(4) \quad \begin{array}{ll} D\llbracket \sim T\lambda \rrbracket_{SD} = 1 & \text{iff “sentence (1) is not true”} \notin \\ & \{o : \text{it’s descending true that } o_D \text{ is descending true in} \\ & \text{English-D}\} \end{array}$$

Now, we can prove the right-hand side of (4). Recall, first, that o_D is the sentence obtained from o when we replace every instance of “true” in English- with “descending true”. Such a replacement would transform (1) into a descending liar, namely a sentence of the form:

$$\sim D\lambda_D,$$

where “ λ_D ” denotes that very sentence. So, the right-hand side of (4) holds just in case it is *not* descending true that λ_D is descending true in English-D. We can thus treat the right-hand side of (4) as equivalent to the following sentence of \mathcal{L} :

$$(5) \quad \sim D\langle D\lambda_D \rangle$$

The following derivation of (5) is ADT-valid:

$$\begin{array}{ll} (i) & | D\langle D\lambda_D \rangle \quad [\text{Hypothesis, for } \textit{reductio}] \\ (ii) & | D\lambda_D \quad [\text{By (i), (D-Out)}] \\ (iii) & | D\langle \sim D\lambda_D \rangle \quad [\text{By (ii), substitution}] \\ (iv) & | \sim D\lambda_D \quad [\text{By (iii), (D-Out)}] \\ (v) & | D\lambda_D \wedge \sim D\lambda_D \quad [\text{By (ii), (iv), conjunction introduction}] \end{array}$$

(vi) $\sim D\langle D\lambda_D \rangle$ [By (i)–(v), *reductio ad absurdum*]

So the right-hand side of (4) holds, and we obtain the result that, contrary to (2b),

(6) ${}^D\llbracket \sim T\lambda \rrbracket_{SD} = 1$.

The postsemantic theory thus yields the surprising result that, from SD, (1) is in fact *descending* true.

This result is troubling. It immediately follows that, from SD, (1) is *safe*—that the liar sentence is not ‘pathological’ after all. This immediately casts Scharp’s account into doubt.

More seriously, the result leads to a contradiction. To see how, notice first that my precise specification of English- was, in certain respects, incidental. I *could* have included “descending true” within English-; after all, I have been using that predicate throughout this paper, and we may wish to understand its semantics. Let us suppose that I did in fact do this, and that I included “D” in ε - to represent the predicate. What axioms should be added to the core semantic theory to reflect these changes? At first glance, one might suggest:

${}^A\llbracket D \rrbracket_s = \{o : \text{it's ascending true that } o \text{ is descending true}\}$

${}^D\llbracket D \rrbracket_s = \{o : \text{it's descending true that } o \text{ is descending true}\}$

However, this is not right: if a sentence o contains the word “true”, then how we evaluate o depends on which alethic standard is operating. Thus, we in fact need the following axioms:

(XI) ${}^A\llbracket D \rrbracket_{sA} = \{o : \text{it's ascending true that } o_A \text{ is descending true in English-A}\}$

(XII) ${}^A\llbracket D \rrbracket_{sD} = \{o : \text{it's ascending true that } o_D \text{ is descending true in English-D}\}$

(XIII) ${}^D\llbracket D \rrbracket_{sA} = \{o : \text{it's descending true that } o_A \text{ is descending true in English-A}\}$

(XIV) ${}^D\llbracket D \rrbracket_{sD} = \{o : \text{it's descending true that } o_D \text{ is descending true in English-D}\}$

Two points should be noted about this. Firstly, by comparing (XII) and (XIV) with (V) and (VII), we can see that, from SD, the semantic values of “D” and “T” are identical. Secondly, none of these amendments affect the result established above that, from SD, sentence (1) is descending true.

Let’s return to that result. From (A-In), it immediately follows that “sentence (1) is descending true” is ascending true from SD. But, on current suppositions, “sentence (1) is descending true” is a sentence of English-, represented in ε - by “ $D\lambda$ ”. So, by the postsemantic theory, it follows that ${}^A\llbracket D\lambda \rrbracket_{SD} = 1$. By the identity of the semantic values of “D” and “T” from SD, it follows that ${}^A\llbracket T\lambda \rrbracket_{SD} = 1$. By (IX), this entails:

(7) ${}^A\llbracket \sim T\lambda \rrbracket_{SD} = 0$.¹⁷

¹⁷ So far as I am aware, there is no direct proof of this, analogous to the argument from (4) to (6) given above. Hence this roundabout way of establishing (7). Thanks to Dave Ripley for pointing me towards this result.

By the postsemantic theory, it follows that “sentence (1) is not true” is *ascending false* from SD.

Pulling the above results together, it follows that, from SD:

(1) is ascending false and (1) is descending true.

This, however, is ruled out by (A-In) and (D-Out): if (1) is ascending false, it is also descending false; and if (1) is descending true, then it is also ascending true. When applied to liar sentences, then, AD semantics yields an inconsistency. As noted above, this directly undermines Scharp’s solution to the liar paradox, AD semantics, and the theoretical motivation for ADT.

4 Negation

These results are deeply puzzling. If ADT is consistent, and AD semantics interprets “T” as expressing ASCENDING TRUTH or DESCENDING TRUTH, then we might naturally expect AD semantics to be consistent. So what is going on?

The problem, I contend, is (IX): Scharp’s recursive rule for negation. It will be helpful to bifurcate the rule as follows:

$$(IX_A) \quad {}^A\llbracket \sim\phi \rrbracket_s = 1 \text{ iff } {}^A\llbracket \phi \rrbracket_s = 0.$$

$$(IX_D) \quad {}^D\llbracket \sim\phi \rrbracket_s = 1 \text{ iff } {}^D\llbracket \phi \rrbracket_s = 0.$$

In the first instance, note that (IX_A) and (IX_D) are disanalogous to ADT: they imply that negating a sentence flips *both* its ascending semantic value (from 1 to 0 and vice versa), *and* its descending semantic value (from 1 to 0 and vice versa). This is not so in ADT generally. Recall:

$$(M) \quad D\langle\phi\rangle \leftrightarrow \sim A\langle\sim\phi\rangle$$

Now, if $\langle\phi\rangle$ is descending true, then it follows from (M) that $\langle\sim\phi\rangle$ is ascending false, and thus by (A-In) and (D-Out) that $\langle\sim\phi\rangle$ is descending false. So negation flips descending truths to descending falsehoods. But the converse does *not* hold. If $\langle\phi\rangle$ is descending false, we cannot automatically infer that $\langle\sim\phi\rangle$ is descending true. We *can* infer that $\langle\sim\phi\rangle$ is *ascending* true, but the descending truth of $\langle\sim\phi\rangle$ only follows from this if $\langle\phi\rangle$ is safe.¹⁸

As ADT is consistent, this immediately suggests that we should tweak (IX_A) and (IX_D) to bring them in line, as follows:

$$(IX_A^*) \quad {}^A\llbracket \sim\phi \rrbracket_s = 1 \text{ iff } {}^D\llbracket \phi \rrbracket_s = 0.$$

$$(IX_D^*) \quad {}^D\llbracket \sim\phi \rrbracket_s = 1 \text{ iff } {}^A\llbracket \phi \rrbracket_s = 0.$$

These axioms, at first glance, better preserve the dual nature of ascending and descending truth encoded by ADT. Certainly, (IX_A^*) directly reflects (M). However, ADT does not, as it stands, include the axiom:

¹⁸ It can also be shown that, if $\langle\phi\rangle$ is ascending false, then $\langle\sim\phi\rangle$ is ascending true. But, if $\langle\phi\rangle$ is ascending true, then we can again only infer that $\langle\sim\phi\rangle$ is ascending false if $\langle\sim\phi\rangle$ is safe.

$$(M^*) \quad \sim D\langle \sim \phi \rangle \leftrightarrow A\langle \phi \rangle.$$

Moreover, although (M^*) is consistent with ADT,¹⁹ so far as I can see it is *not* a theorem of ADT.²⁰ If this is right, then (IX_D^*) does not directly reflect ADT. If Scharp intends AD semantics to cohere with ADT, he will ultimately need to include (M^*) as an axiom of ADT.

Regardless, there are two good reasons to prefer (IX_A^*) and (IX_D^*) to (IX) . The first is formal. Suppose we construct an AD semantic theory for some fragment of English, using the negation clauses (IX_A^*) and (IX_D^*) . Assume that, when two sentences in the fragment differ only in that one is the negation of the other, the presemantic theory differs in their representation by including “ \sim ” at the front of the former. Given the postsemantic theory, we obtain the following from (IX_A^*) to (IX_D^*) :

The negated sentence is ascending/descending true from s iff the un-negated sentence is not descending/ascending true from s .

Assuming a fixed alethic standard, we can easily reformulate these in \mathcal{L} :

$$\begin{aligned} A\langle \sim \phi \rangle &\leftrightarrow \sim D\langle \phi \rangle \\ D\langle \sim \phi \rangle &\leftrightarrow \sim A\langle \phi \rangle \end{aligned}$$

which, of course, are equivalent to (M) and (M^*) respectively. Thus, as (M) and (M^*) are consistent with ADT, we can see that (IX_A^*) and (IX_D^*) are unproblematic.

The second reason to prefer (IX_A^*) and (IX_D^*) is philosophical. Suppose that $\langle \phi \rangle$ is unsafe: it is ascending true and descending false. Then, what should we think of $\langle \sim \phi \rangle$? At face value, we might expect the negation of an unsafe sentence to *also* be unsafe: negating a ‘pathological’ sentence does not negate its ‘pathological’ nature. This line of thought is borne out by (IX_A^*) and (IX_D^*) : if $\langle \phi \rangle$ is unsafe (i.e. $^A\llbracket \phi \rrbracket_s = 1$ and $^D\llbracket \phi \rrbracket_s = 0$, then they entail that $\langle \sim \phi \rangle$ is *also* unsafe (i.e. $^A\llbracket \sim \phi \rrbracket_s = 1$ and $^D\llbracket \sim \phi \rrbracket_s = 0$).

This is in stark contrast to (IX_A) and (IX_D) . These axioms entail that, if $\langle \phi \rangle$ is unsafe, then $\langle \sim \phi \rangle$ is ascending false and descending true (i.e. $^A\llbracket \sim \phi \rrbracket_s = 0$ and

¹⁹ Bacon’s model of ADT straightforwardly satisfies (M^*) as a direct consequence of Bacon’s definition of “ A ”. So we can consistently add (M^*) to ADT. See Bacon (2019: 381f).

²⁰ An anonymous reviewer notes that they were initially misled by the following derivation:

- (i) $D\langle \sim \phi \rangle \leftrightarrow \sim A\langle \sim \sim \phi \rangle$ [by (M)]
- (ii) $D\langle \sim \phi \rangle \leftrightarrow \sim A\langle \phi \rangle$ [by $*$]
- (iii) $\sim D\langle \sim \phi \rangle \leftrightarrow A\langle \phi \rangle$ [logic]

However, as they note, the step from (i) to (ii) appears to be invalid in ADT. The following principle would justify the step:

$$(*) \quad (\phi \leftrightarrow \psi) \leftrightarrow (A\langle \phi \rangle \leftrightarrow A\langle \psi \rangle).$$

But, the reviewer points out, $(*)$ is not valid according to ADT: just let ϕ be $D\lambda_D$ and let ψ by $1=0$. Even if we require ϕ and ψ to be logically equivalent, it is unclear that the resulting principle would hold in ADT: Scharp writes that it is an “essential feature of ADT” that “logically interdeducible sentences might have different descending truth values or different ascending truth values” (2013a: 173).

$^D\llbracket \sim\phi \rrbracket_s = 1$)—an impossibility. This is precisely the problem that arose when we applied AD semantics to (1): the sentence “ $T\lambda$ ” is unsafe, so the rules for negation entail that “ $\sim T\lambda$ ” is ascending false and descending true.

Let us define *AD* semantics* to be an approach to the empirical study of meaning in natural language, differing from AD semantics only insofar as, whereas the latter governs negation with (IX), the former governs negation with (IX_A*) and (IX_D*). If we apply AD* semantics to (1), the presemantic theory will again represent (1) with $\sim T\lambda$. We can then derive the following:

- (8) (a) $^A\llbracket \sim T\lambda \rrbracket_{SD} = 1$ iff $^D\llbracket T\lambda \rrbracket_{SD} = 0$ [By (IX_A*)]
 (b) $^D\llbracket \sim T\lambda \rrbracket_{SD} = 1$ iff $^A\llbracket T\lambda \rrbracket_{SD} = 0$ [By (IX_D*)]

and subsequently:

- (9) (a) $^A\llbracket \sim T\lambda \rrbracket_{SD} = 1$ iff $^D\llbracket \lambda \rrbracket_{SD} \notin ^D\llbracket T \rrbracket_{SD}$ [By (8a) and (VIII)]
 (b) $^D\llbracket \sim T\lambda \rrbracket_{SD} = 1$ iff $^A\llbracket \lambda \rrbracket_{SD} \notin ^A\llbracket T \rrbracket_{SD}$ [By (8b) and (VIII)]

The right-hand side of (9a) is equivalent to the right-hand side of (4), which we proved above. So it follows from (9a) that $^A\llbracket \sim T\lambda \rrbracket_{SD} = 1$, and thus from the post-semantic theory that “sentence (1) is not true” is ascending true from SD.

Moreover, if we add (M*) to ADT as I have recommended, then we can show that the right-hand side of (9b) is false. We note that the right-hand side of (9b) is equivalent to the following sentence of \mathcal{L} :

$$\sim A\langle D\lambda_D \rangle,$$

which is easily shown to be false. Thus, it follows that $^D\llbracket \sim T\lambda \rrbracket_{SD} = 0$ and, by the postsemantic theory, that “sentence (1) is not true” is descending false from SD.

From SD, then, it can be shown using AD* semantics that (1) is ascending true and descending false, and that (1) is thus unsafe. We obtain the same result from SA. When applied to the liar sentence, AD* semantics recovers the intuitively right result.

5 Revenge

AD* semantics also gives Scharp the resources to respond to a revenge paradox, which I introduced in Pinder (2015). I begin by reconstructing the revenge paradox.

The guiding idea is that, given AD semantics, one might ask of a given sentence whether it is *true-from-some-context*. Consider the sentence:

- (10) Sentence (10) is not true-from-some-context.

Let us add “sentence (10)” and “true-from-some-context” to English-, and suppose that our presemantic theory represents them in ε - with “ ρ ” and “ C ” respectively. Thus, (10) is represented:

$$\sim C\rho.$$

We then add the following rules to our core semantic theory²¹:

- (XV) $\mathcal{V}\llbracket \rho \rrbracket_s = \text{"sentence (10) is not true-from-some-context"}$.
- (XVI) $\mathcal{A}\llbracket C \rrbracket_s = \{o : \text{it's ascending true that } o_A \text{ is ascending true from SA, or it's ascending true that } o_D \text{ is descending true from SD}\}$
- (XVII) $\mathcal{D}\llbracket C \rrbracket_s = \{o : \text{it's descending true that } o_A \text{ is ascending true from SA, or it's descending true that } o_D \text{ is descending true from SD}\}$

From here, using AD semantics, we can derive the following biconditionals:

- (11) (a) $\mathcal{A}\llbracket \sim C \rho \rrbracket_s = 1$ iff $\mathcal{A}\llbracket \rho \rrbracket_s \notin \mathcal{A}\llbracket C \rrbracket_s$
- (b) $\mathcal{D}\llbracket \sim C \rho \rrbracket_s = 1$ iff $\mathcal{D}\llbracket \rho \rrbracket_s \notin \mathcal{D}\llbracket C \rrbracket_s$

Notice that “sentence (10) is not true-from-some-context” does not contain a truth predicate, so using (XV)–(XVII), we can infer:

- $\mathcal{A}\llbracket \sim C \rho \rrbracket_s = 1$ iff it's not ascending true that “sentence (10) is not true-from-some-context” is ascending true from SA, and it's not ascending true that “sentence (10) is not true-from-some-context” is descending true from SD
- $\mathcal{D}\llbracket \sim C \rho \rrbracket_s = 1$ iff it's not descending true that “sentence (10) is not true-from-some-context” is ascending true from SA, and it's not descending true that “sentence (10) is not true-from-some-context” is descending true from SD

By the postsemantic theory, we obtain the following ascending/descending truth conditions:

“Sentence (10) is not true-from-some-context” is ascending true from s iff
it's not ascending true that “sentence (10) is not true-from-some-context” is ascending true from SA, and it's not ascending true that “sentence (10) is not true- from-some-context” is descending true from SD.

“Sentence (10) is not true-from-some-context” is descending true from s iff
it's not descending true that “sentence (10) is not true-from-some-context” is ascending true from SA, and it's not descending true that “sentence (10) is not true- from-some-context” is descending true from SD.

We can translate these ascending/descending truth conditions into ADT as follows:

$$\begin{aligned} \mathcal{A}\langle \sim C \rho \rangle_{[s]} &\leftrightarrow \sim \mathcal{A}\langle \mathcal{A}\langle \sim C \rho \rangle \rangle_{[SA]} \wedge \sim \mathcal{A}\langle \mathcal{D}\langle \sim C \rho \rangle \rangle_{[SD]} \\ \mathcal{D}\langle \sim C \rho \rangle_{[s]} &\leftrightarrow \sim \mathcal{D}\langle \mathcal{A}\langle \sim C \rho \rangle \rangle_{[SA]} \wedge \sim \mathcal{D}\langle \mathcal{D}\langle \sim C \rho \rangle \rangle_{[SD]} \end{aligned}$$

²¹ I change the notation from Pinder (2015), bringing it in line with the notation herein.

where “ p ” is taken to be intersubstitutable with “ $\langle \sim Cp \rangle$ ” and the subscripts keep track of the alethic standard in operation.²² And, finally, simplifying these sentences by substituting “ p ” for “ $\langle \sim Cp \rangle$ ”, we obtain²³:

- (12) (a) $Ap_{[\cdot]} \leftrightarrow \sim A\langle Ap \rangle_{[SA]} \wedge \sim A\langle Dp \rangle_{[SD]}$
 (b) $Dp_{[\cdot]} \leftrightarrow \sim D\langle Ap \rangle_{[SA]} \wedge \sim D\langle Dp \rangle_{[SD]}$

We need two additional results to derive the revenge paradox. The first is obtained with reasoning analogous to the above (which I omit for brevity):

- (13) (a) $A\langle Cp \rangle_{[\cdot]} \leftrightarrow A\langle Ap \rangle_{[SA]} \vee A\langle Dp \rangle_{[SD]}$
 (b) $D\langle Cp \rangle_{[\cdot]} \leftrightarrow D\langle Ap \rangle_{[SA]} \vee D\langle Dp \rangle_{[SD]}$

The second is a corollary of (12): the ascending and descending truth values of the revenge sentence are context of assessment invariant. That is:

- (14) (a) $Ap_{[SA]} \leftrightarrow Ap_{[SD]}$
 (b) $Dp_{[SA]} \leftrightarrow Dp_{[SD]}$

Using these results, we can derive a contradiction.²⁴ We first derive that p is not true-from-some-context:

- | | | |
|--------|--|---|
| (i) | $ Ap_{[SA]}$ | [Hypothesis, for <i>reductio</i>] |
| (ii) | $ \sim A\langle Ap \rangle_{[SA]}$ | [By (i), (12a \rightarrow), \wedge -elimination] |
| (iii) | $ \sim Ap_{[SA]}$ | [By (ii), (A-In), <i>modus tollens</i>] |
| (iv) | $\sim Ap_{[SA]}$ | [By (i)–(iii), <i>reductio ad absurdum</i>] |
| (v) | $\sim Dp_{[SA]}$ | [By (iv), (A-In), (D-Out), <i>modus tollens</i>] |
| (vi) | $\sim Dp_{[SD]}$ | [By (v), (14b)] |
| (vii) | $\sim Ap_{[SA]} \wedge \sim Dp_{[SD]}$ | [By (iv), (vi), \wedge -introduction] |
| (viii) | $\sim (Ap_{[SA]} \vee Dp_{[SD]})$ | |

From here, we derive that p is true-from-some-context:

- | | | |
|-------|--|---|
| (ix) | $\sim A\langle \sim Cp \rangle_{[SA]}$ | [By (iv), substitution] |
| (x) | $D\langle Cp \rangle_{[SA]}$ | [By (ix), (M)] |
| (xi) | $D\langle Ap \rangle_{[SA]} \vee D\langle Dp \rangle_{[SD]}$ | [By (x), (13b \rightarrow)] |
| (xii) | $Ap_{[SA]} \vee Dp_{[SD]}$ | [By (xi), (D-Out), disjunctive reasoning] |

And, finally:

- (xiii) \perp [By (viii) and (xii)]

Thus, AD semantics admits of a *revenge* paradox.

²² I assume here that, given Scharp’s comments about his semantics for “true”, these subscripts should apply to whole terms, and should not iterate. Thus, using (D-Out), we can infer “ $Dp_{[SD]}$ ” from “ $D\langle Dp \rangle_{[SD]}$ ”; and sentences such as “ $D(Dp_{[SA]})_{[SD]}$ ” are not well-formed.

²³ This substitution is explicitly permitted. See Scharp (2013a: 154).

²⁴ I have simplified the derivation from Pinder (2015) slightly.

Aimed at AD* semantics, however, this derivation does not get off the ground. When deriving the ascending and descending truth conditions for ρ , we do not derive (11), but rather:

$$\begin{aligned} {}^A\llbracket \sim C\rho \rrbracket_s &= 1 \text{ iff } {}^D\llbracket \rho \rrbracket_s \notin {}^D\llbracket C \rrbracket_s \\ {}^D\llbracket \sim C\rho \rrbracket_s &= 1 \text{ iff } {}^A\llbracket \rho \rrbracket_s \notin {}^A\llbracket C \rrbracket_s \end{aligned}$$

And thus, rather than (12), we obtain:

$$\begin{aligned} (15) \quad (a) \quad A\rho_{[.]} &\leftrightarrow \sim D\langle A\rho \rangle_{[SA]} \wedge \sim D\langle D\rho \rangle_{[SD]} \\ (b) \quad D\rho_{[.]} &\leftrightarrow \sim A\langle A\rho \rangle_{[SA]} \wedge \sim A\langle D\rho \rangle_{[SD]} \end{aligned}$$

As (12a) plays a pivotal role in the derivation of the contradiction—at step (ii)—this blocks my derivation of this revenge paradox. Moreover, we can use (15a, b) to prove that ρ is in fact *unsafe*. First, we can prove that ρ is ascending true from SA.

(i)	$\sim A\rho_{[SA]}$	[Hypothesis, for <i>reductio</i>]
(ii)	$D\langle A\rho \rangle_{[SA]} \vee D\langle D\rho \rangle_{[SD]}$	[By (i), (15a)]
(iii)	$D\langle A\rho \rangle_{[SA]}$	[Hypothesis, for <i>reductio</i>]
(iv)	$A\rho_{[SA]}$	[By (iii), (D-Out)]
(v)	$A\rho_{[SA]} \wedge \sim A\rho_{[SA]}$	[By (i), (iv)]
(vi)	$\sim D\langle A\rho \rangle_{[SA]}$	[By (iii)–(v), <i>reductio ad absurdum</i>]
(vii)	$D\langle D\rho \rangle_{[SD]}$	[By (ii), (vi), disjunctive syllogism]
(viii)	$D\rho_{[SD]}$	[By (vii), (D-Out)]
(ix)	$\sim A\langle A\rho \rangle_{[SA]} \wedge \sim A\langle D\rho \rangle_{[SD]}$	[By (viii), (15b)]
(x)	$\sim A\langle D\rho \rangle_{[SD]}$	[By (ix)]
(xi)	$\sim D\rho_{[SD]}$	[By (x), (A-In), <i>modus tollens</i>]
(xii)	$D\rho_{[SD]} \wedge \sim D\rho_{[SD]}$	[By (viii), (xi)]
(xiii)	$A\rho_{[SA]}$	[By (i)–(xii), <i>reductio ad absurdum</i>]

So ρ is ascending true from SA. Second, we can prove that ρ is descending false from SA.

(xiv)	$D\rho_{[SA]}$	[Hypothesis, for <i>reductio</i>]
(xv)	$\sim A\langle A\rho \rangle_{[SA]} \wedge \sim A\langle D\rho \rangle_{[SD]}$	[By (xiv), (15b)]
(xvi)	$\sim A\langle A\rho \rangle_{[SA]}$	[By (xv)]
(xvii)	$\sim A\rho_{[SA]}$	[By (xvi), (A-In), <i>modus tollens</i>]
(xviii)	$D\rho_{[SD]} \wedge \sim D\rho_{[SD]}$	[By (xvii), and (i)–(xii) above]
(xix)	$\sim D\rho_{[SA]}$	[By (xiv)–(xviii), <i>reductio ad absurdum</i>]

So ρ is descending false from SA. Finally, given that (15a, b) imply that ρ is context of assessment invariant, it also follows that ρ is ascending true and descending false

from SD. The upshot is that p is unsafe from both contexts of assessment—intuitively the right result.

All of this suggests that the revenge paradox presented in Pinder (2015) is, in fact, another symptom of AD semantics' mishandling of negation. When we shift to AD* semantics, however, the problem appears to be resolved. As things stand, no revenge paradox has even been raised against AD* semantics.

6 Adequacy

Is AD* semantics consistent? Given that AD* semantics is an approach to natural language semantics, partially constituted by *informal* theories, the question is not well-posed. The threat of inconsistency only arises when we take AD* semantics to have consequences that we can reformulate in \mathcal{L} , the language of ADT.

The basic threat derives from the axioms of the core semantic theory. On minimal assumptions about the presemantic and postsemantic theories, the core semantic theory generates substantive results that we can reformulate in \mathcal{L} . When we do this, there are two important threats to distinguish. First, we might find that the recursive clauses governing the connectives lead to results that are inconsistent with the axioms of ADT. This is essentially the problem that I highlighted in Sect. 4: the recursive clause (IX), when applied to unsafe sentences, yields results that, formulated in \mathcal{L} , are inconsistent with (A-In) and (D-Out). Second, we might invent a new term that extends natural language such that, when we formulate corresponding axioms in the core semantic theory, we can obtain a theorem that, formulated in \mathcal{L} , is inconsistent with ADT. This is essentially the problem I highlighted in Sect. 5: in Pinder (2015), I introduced the term “true-from-some-context”, motivating the axioms (XV)–(XVII), leading to ascending/descending truth conditions that, reformulated in \mathcal{L} , lead to contradiction.

Neither threat of inconsistency can be ruled out here. Firstly, it is unclear that the possibility of a revenge paradox could ever be ruled out. One might show that AD* semantics, when applied to *some* fragments of language, only yields results compatible with ADT. But, as AD* semantics is an approach to the empirical study of meaning in natural languages, and natural languages can always be extended with novel terms and concepts, there always remains the possibility of revenge. If we invent or discover a novel concept (such as TRUTH-FROM-SOME-CONTEXT), we are justified in formulating new axioms in our core semantic theory to govern a term that expresses that concept. If we can reverse engineer a concept to justify axioms that yield results that, when formulated in \mathcal{L} , are inconsistent with ADT, we will find a revenge paradox. I do not know how to prove that such a project must be unsuccessful.

More interestingly, I cannot show that the recursive clauses governing connectives are unproblematic, because the recursive definition *outstrips* ADT. I have already suggested that the second negation clause (IX_D*) outstrips ADT as it stands, as the latter appears not to include (M*) amongst its theorems. Another example is conjunction. Bifurcating (X) for ease, the core semantic theory gives us:

$$(X_A) \quad \text{A}[\phi \wedge \psi]_s = 1 \text{ iff } \text{A}[\phi]_s = 1 \text{ and } \text{A}[\psi]_s = 1.$$

$$(X_D) \quad \text{D}[\phi \wedge \psi]_s = 1 \text{ iff } \text{D}[\phi]_s = 1 \text{ and } \text{D}[\psi]_s = 1.$$

In contrast, recall that the relevant axioms of ADT are:

$$(\wedge\text{-A}) \quad (\text{A}(\phi) \wedge \text{A}(\psi)) \rightarrow \text{A}(\phi \wedge \psi)$$

$$(\wedge\text{-D}) \quad \text{D}(\phi \wedge \psi) \rightarrow (\text{D}(\phi) \wedge \text{D}(\psi))$$

So, AD* semantics gives us *biconditionals*, only one direction of which are given by $(\wedge\text{-A})$ and $(\wedge\text{-D})$. Moreover, the converses of $(\wedge\text{-A})$ and $(\wedge\text{-D})$ are *not* valid in ADT.²⁵ The same point applies mutatis mutandis to disjunction.²⁶ And, while the core semantic theory governs the conditional thus:

$$\text{V}[\phi \rightarrow \psi]_s = 1 \text{ iff if } \text{V}[\phi]_s = 1 \text{ then } \text{V}[\psi]_s = 1$$

there is no corresponding rule in ADT governing the conditional at all.²⁷ So the recursive clauses governing the connectives in the core semantic theory significantly outstrip ADT.

What this shows is that AD* semantics, when combined with ADT, is much stronger than ADT alone. The immediate consequence is that, even putting revenge paradoxes aside, the consistency of ADT does not suffice to protect AD* semantics from the threat of inconsistency. Although AD* semantics seems to deal adequately with liar sentences such as (1), it remains possible that there are alethic paradoxes that AD* semantics cannot handle.

This raises an important difficulty for Scharp. The concepts *ASCENDING TRUTH* and *DESCENDING TRUTH* have been designed to play the central theoretical role in AD* semantics. What the present discussion shows is that those concepts are not, as they stand, adequate for their intended task. Their role in AD* semantics demands more than is supplied by ADT. So Scharp needs *either* to strengthen ADT, *or* to weaken AD* semantics. The former option requires care, and there is no guarantee of success. As Bacon (2019: 383–385) has discussed, there are several ways that Scharp might seek to strengthen ADT, but the options are not all jointly compatible. It remains an open question whether Scharp can strengthen ADT sufficiently to support its role in AD* semantics, without leading to inconsistency. In contrast, it is unclear whether the latter option is even viable. For example, the additional strength of AD* semantics derives in part from the fact that, where ADT has conditionals, AD* semantics

²⁵ In Bacon's model of ADT, the interpretation of "D" is not closed under conjunction introduction. So we obtain a counterexample to the converse of $(\wedge\text{-D})$ by letting ϕ and ψ be different axioms of ADT; and to the converse of $(\wedge\text{-A})$ by letting ϕ and ψ be the negations of different axioms of ADT. See Bacon (2019: 381–382, 386).

²⁶ See Scharp (2013a: 154, 248).

²⁷ As it stands, this axiom *does* lead to results that are inconsistent with ADT. This can be demonstrated by considering a sentence of the form:

(16) If sentence (16) is descending true, then sentence (16) is not descending true.

(I omit the details.) As Scharp does not discuss conditionals, however, I assume that this constitutes future research, and so I will not push the point here.

has biconditionals. But having biconditionals is non-negotiable: if AD* semantics is to be a viable approach to natural language semantics, it must yield conditions necessary and sufficient for the ascending/descending truth of each sentence in the relevant fragment of natural language. Such conditions are expressed with biconditionals.

The upshot is this. Even if AD* semantics can deal adequately with the liar paradox and my earlier revenge paradox, it remains unclear whether AD* semantics deals adequately with alethic paradoxes more generally. The underlying problem is that the concepts ASCENDING TRUTH and DESCENDING TRUTH are not, as they stand, adequate to play the central theoretical role in AD* semantics—and that it is unclear whether those concepts can be further developed to make them adequate for this theoretical role. As things stand, the consistency of ASCENDING TRUTH and DESCENDING TRUTH does not guarantee that AD* semantics can be consistently applied to ‘pathological’ sentences in general.²⁸

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bacon, A. (2019). Sharp on replacing truth. *Inquiry*, 62(4), 370–386.
- Pinder, M. (2015). A revenge problem without the concept of truth. *Thought: A Journal of Philosophy*, 4(3), 151–161.
- Predelli, S. (2005). *Contexts: Meaning, truth, and the use of language*. Oxford: Oxford University Press.
- Sharp, K. (2007). Replacing truth. *Inquiry*, 50(6), 606–621.
- Sharp, K. (2013a). *Replacing truth*. Oxford: Oxford University Press.
- Sharp, K. (2013b). Truth, the liar, and relativism. *Philosophical Review*, 122(3), 427–510.
- Sharp, K. (2014). Truth, revenge, and internalizability. *Erkenntnis*, 79(3), 597–645.
- Sharp, K. (2019). Replies to Bacon, Eklund, and Greenough on replacing truth. *Inquiry*, 62(4), 422–475.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

²⁸ An earlier version of this paper was presented at the *Semantic Paradox and Revenge Workshop* at the University of Salzburg, and the paper improved immeasurably as a result of the subsequent discussion. I thank Dave Ripley and Andrew Bacon in particular for their comments and suggestions. I am also grateful to two anonymous referees for this journal, whose insightful comments pushed me to develop and improve the paper even further.